

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2001, 15:28:47 ; Search time 2667.65 Seconds

(without alignments)
2626.615 Million cell updates/sec

Title: US-09-784-340-3_COPY_18322_18774

Perfect score: 453
Sequence: 1 gtaagtactactgtctgttac.....tgtgtgttttccctccacg 453

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: gb_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_rod:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_pr1:*
95: gb_pr2:*
96: gb_pr3:*
97: gb_pr4:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	% Match	Query Length	DB ID	Description
1	451.4	99.6	212904	66	AC021146	AC021146 Homo sapi
2	359	79.2	138665	72	AC055794	AC055794 Homo sapi
3	289	63.8	402	54	G50188	G50188 SHGC-79709
4	48.2	10.6	894	53	CNS06KRO	AL403234 T3 end of
5	48	10.6	161380	62	AC011683	AC011683 Homo sapi
6	48	10.6	163233	80	AL356974	AL356974 Homo sapi
7	48	10.6	182162	69	AC025302	AC025302 Homo sapi
8	47.6	10.5	177341	68	AC023596	AC023596 Homo sapi

C	9	47.2	10.4	139104	67	AC022772	Homo sapi
C	10	47	10.4	1355	12	AF031433	Gossypium
C	11	45.8	10.1	162763	90	AL359262	Human DNA
C	12	45.8	10.1	170274	64	AC016462	Human DNA
C	13	45.8	10.1	186695	63	AC013504	Human DNA
C	14	45.6	10.1	182232	79	AL353783	Human DNA
C	15	45.4	10.0	169246	62	AC011254	Human DNA
C	16	45	9.9	1137	53	CNS071NS	Human DNA
C	17	45	9.9	158250	82	AL590243	Human DNA
C	18	45	9.9	169437	81	AL391538	Human DNA
C	19	44.8	9.9	85382	90	AL390793	Human DNA
C	20	44.8	9.9	123593	80	AL356912	Human DNA
C	21	44.8	9.9	14586	90	AL354831	Human DNA
C	22	44.8	9.9	155813	65	AC019350	Human DNA
C	23	44.8	9.9	178433	81	AL450447	Human DNA
C	24	44.8	9.9	194100	65	AC018674	Human DNA
C	25	44.6	9.8	100269	13	ATF18022	Arabidops
C	26	44.6	9.8	135350	81	AL391626	Human DNA
C	27	44.6	9.8	164689	80	AL359928	Human DNA
C	28	44.4	9.8	7218	10	166494	Sequence 14
C	29	44.4	9.8	171300	61	AC010941	Human DNA
C	30	44.4	9.8	186221	62	AC012530	Human DNA
C	31	44	9.7	11869	5	AF298624	Diclyoste
C	32	44	9.7	170866	80	AL357043	Human DNA
C	33	44	9.7	176314	79	AL158044	Human DNA
C	34	44	9.7	192128	63	AC013382	Human DNA
C	35	43.8	9.7	3630	14	MISC14	Human DNA
C	36	43.8	9.7	6340	15	YSCMCG03	Yeast mitoc
C	37	43.8	9.7	138232	63	AC012582	Human DNA
C	38	43.6	9.6	158250	66	AC020781	Human DNA
C	39	43.6	9.6	196721	87	AC019179	Human DNA
C	40	43.4	9.6	111554	13	AP002460	Arabidops
C	41	43.4	9.6	175986	67	AC023071	Human DNA
C	42	43.4	9.6	176666	83	CNS01RG6	Human DNA
C	43	43.2	9.5	179947	66	AC021467	Human DNA
C	44	43.2	9.5	48128	9	AX059517	Sequence
C	45	43.2	9.5	94503	12	AC007730	Arabidops

ALIGNMENTS

RESULT	1	AC021146	212904 bp	DNA	07-JUL-2000
LOCUS		AC021146	212904 bp	DNA	07-JUL-2000
DEFINITION		Human sapiens chromosome 4 clone RP11-468N14, WORKING DRAFT			
ACCESSION		AC021146			
VERSION		AC021146.4	GT:8568861		
KEYWORDS		HTG; HTGS_PHASE1; HTGS_DRAFT.			
SOURCE		human.			
ORGANISM		Homo sapiens			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.			
JOURNAL		Waterston, R.H.			
REFERENCE		1 (bases 1 to 212904)			
TITLE		The sequence of Homo sapiens clone			
AUTHORS		Waterston, R.H.			
JOURNAL		Unpublished			
REFERENCE		2 (bases 1 to 212904)			
TITLE		Direct Submission			
JOURNAL		Submitted (14-JAN-2000) Genome Sequencing Center, Washington			
COMMENT		University School of Medicine, 4444 Forest Park Parkway, St. Louis,			
		MO 63108, USA			
		On Jun 16, 2000 this sequence version replaced g1:7344259.			

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H.NH0468N14
----- Summary Statistics -----

FEATURES

Source

Sequencing vector: M13: 88%
Sequencing vector: plasmid: 12%
Chemistry: Dye-terminator Big Dye, 12% of reads
Chemistry: Dye-terminator Big Dye, 12% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 199729 bases at least Q40
Consensus quality: 203731 bases at least Q30
Consensus quality: 206340 bases at least Q20
Insert size: 213000; agarose-gel
Insert size: 210604; sum-of-contigs
Quality coverage: 3.75 in Q20 bases; agarose-gel
Quality coverage: 3.83 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1564: contig of 1564 bp in length
1565: gap of unknown length
1664: contig of 1421 bp in length
3085: gap of unknown length
3086: gap of unknown length
3185: contig of 1949 bp in length
5134: gap of unknown length
5135: gap of unknown length
5235: contig of 2835 bp in length
8169: gap of unknown length
8269: gap of unknown length
8170: contig of 2072 bp in length
10341: gap of unknown length
10342: contig of 3173 bp in length
10442: gap of unknown length
13614: contig of 3375 bp in length
13615: gap of unknown length
13715: contig of 3375 bp in length
13716: gap of unknown length
17189: contig of 3512 bp in length
17190: gap of unknown length
20701: contig of 3512 bp in length
20801: gap of unknown length
20802: contig of 4200 bp in length
25001: gap of unknown length
25002: contig of 3919 bp in length
25102: gap of unknown length
29120: contig of 4236 bp in length
29121: gap of unknown length
33357: gap of unknown length
33457: contig of 3446 bp in length
36902: gap of unknown length
37002: contig of 5973 bp in length
42975: gap of unknown length
43075: gap of unknown length
48339: contig of 5264 bp in length
48340: gap of unknown length
48440: contig of 8090 bp in length
56529: gap of unknown length
56530: gap of unknown length
56513: contig of 8584 bp in length
65214: gap of unknown length
65314: contig of 9402 bp in length
74715: gap of unknown length
74815: contig of 13731 bp in length
74816: gap of unknown length
88547: gap of unknown length
88548: contig of 14721 bp in length
103367: gap of unknown length
103368: gap of unknown length
103467: contig of 14700 bp in length
103468: gap of unknown length
118167: gap of unknown length
118168: contig of 14498 bp in length
132765: gap of unknown length
132766: contig of 21227 bp in length
132866: gap of unknown length
132867: contig of 21227 bp in length
154192: gap of unknown length
154193: contig of 27952 bp in length
182144: gap of unknown length
182145: contig of 30660 bp in length
182245: gap of unknown length
212904: contig of 30660 bp in length.

1. 212904
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"

misc-feature	140047..145961	/note="assembly-fragment:00462 fragment_chain:2"
misc-feature	146062..163233	/note="assembly-fragment:00430 fragment_chain:2"
BASE COUNT	54411 a 28101 c 29009 g 51305 t	407 others
ORIGIN		
Query Match	10.6%; Score 48; DB 80; Length 163233;	
Best Local Similarity	50.9%; Pred. NO. 2.3;	
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;		
Oy	96 gatttactccaacatattgagctatcattcaacccttggtaagataactcttca	155
Db 114443	gTTTTTTGTAGACACTAATGATGATGCTTTTGTTCATTTCCCAATGTTGTTTAATT	114384
Oy	156 gtagctacagagtgcaaccttcttcgaatataggttttaagtaaacctgtgc	215
Db 114383	GTAGTGTTTAATACATTTTAATTTAATATAGTACTGATAAGTTTGAATTCACATCTATT	114324
Oy	216 ttaccaagctttattacatcattatattaccacatttggtaagataactcttca	275
Db 114323	ATTTTGGCATTTGGCTTTACACATTTGATATATCTATCTATCTATTTCCATTTCTTTTA	114264
Oy	276 gtcttcacatactatcgtttaatactatgtaccacaataat 319	
Db 114263	TTTTGTTCAGTACAGATGTTTATATATATAGCCATTTAATTACT	114220
RESULT 7		
AC025302	AC025302 182162 bp DNA HTG 25-MAY-2000	
LOCUS	Homo sapiens clone RP11-680K19, WORKING DRAFT SEQUENCE, 14	
DEFINITION	unordered pieces.	
ACCESSION	AC025302	
VERSION	AC025302.3 GI:8072561	
KEYWORDS	HTG; HTGS-PHASE1; HTGS-DRAFT.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 182162)	
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beka,F., Bonuslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A., Choepey,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearbellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gaidyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kamp,L., Karatas,A., Klein,J., Labrocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurt,A., McKernan,K., McPheters,R., Melidiri,T., Meneses,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanli,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Seery,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J., Vassiliadis,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.	
TITLE	Direct Subassembly	
JOURNAL	Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
COMMENT	On May 25, 2000 this sequence version replaced gi:17280312.	

```

All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4330
Center clone name: 680_K.19
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 176346 bases at least Q40
Consensus quality: 179048 bases at least Q30
Consensus quality: 180129 bases at least Q20
Insert size: 183000; agarose-fp
Insert size: 180862; sum-of-contigs
Quality coverage: 5.0 in Q20 bases; agarose-fp
Quality coverage: 5.1 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1322: contig of 1322 bp in length
*
* 1323 1422: gap of 100 bp
*
* 1423 4198: contig of 2776 bp in length
*
* 4199 4298: gap of 100 bp
*
* 4299 6627: contig of 2329 bp in length
*
* 6628 6727: gap of 100 bp
*
* 6728 9902: contig of 3175 bp in length
*
* 9903 10002: gap of 100 bp
*
* 10003 13206: contig of 3204 bp in length
*
* 13207 13306: gap of 100 bp
*
* 13307 17106: contig of 3800 bp in length
*
* 17107 17206: gap of 100 bp
*
* 17207 21975: contig of 4769 bp in length
*
* 21976 22075: gap of 100 bp
*
* 22076 24992: contig of 2917 bp in length
*
* 24993 25092: gap of 100 bp
*
* 25093 33438: contig of 8346 bp in length
*
* 33439 33538: gap of 100 bp
*
* 33539 49399: contig of 15861 bp in length
*
* 49400 49499: gap of 100 bp
*
* 49500 69413: contig of 19914 bp in length
*
* 69414 69513: gap of 100 bp
*
* 69514 91950: contig of 22437 bp in length
*
* 91951 92050: gap of 100 bp
*
* 92051 119326: contig of 27276 bp in length
*
* 119327 119426: gap of 100 bp
*
* 119427 182162: contig of 62736 bp in length.
*
* Location/Qualifiers
*
* 1. 182162
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /clone="RP11-680K19"
* /clone_11b="RPCT-11 Human Male BAC"
*
* 1. 1322
* /note="assembly_fragment"
*
* 1423. 4198
* /note="assembly_fragment"
*
* 4299. 6627
* /note="assembly_fragment"
*
* 6728. 9902
* /note="assembly_fragment"
*
* 10003. 13206

```

```

misc_feature /note="assembly-fragment"
13307..17106
/note="assembly-fragment"
17207..21975
/note="assembly-fragment"
22076..24992
/note="assembly-fragment"
clone_end:17
vector_side:left"
25093..33438
/note="assembly-fragment"
33539..49399
/note="assembly-fragment"
clone_end:56
vector_side:left"
49500..69413
/note="assembly-fragment"
69514..91950
/note="assembly-fragment"
92051..119326
/note="assembly-fragment"
119427..182162
/note="assembly-fragment"
misc_feature /note="assembly-fragment"
58607 a 31962 c 32284 g 58008 t 1301 others
ORIGIN
Query Match 10.6% Score 48; DB 69; Length 182162;
Best Local Similarity .50.9%; Pred. No. 2.2;
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 96 gatttaccatccatgagtcacatcctgtcactgacgaatgttggaaatt 155
16544 GTTATTTGAGACAGTAAGATGATGCTTTGTTGTCATTCGCCAATGTTTAAAT 16603
156 gtatccatagagtgcaactctcttcacatgaatattaggttaagtaacaaatgac 215
16604 GTAGCTTTTAATCATTTAAATTAATATAGTACTGATGAATGTTAGATTCACATCTAAT 16663
QY 216 ttactaagctttattacacatcttaattaccacatttggtaagaatactcttca 275
16664 ATTTGCTATTGCTTTCACACATTTGATATATCTATCTATTTCTCCATTTCTCTTTA 16723
QY 276 gtcttcacatatactgtttactactgttaacacaaatatt 319
16724 TTTTGTCTCAGTAGATGTTTATATATATGCCATTTAAATTAATTA 16767
DB AC023596 177341 bp DNA HTG 04-NOV-2000
LOCUS Homo sapiens chromosome 3 clone RP11-699J21, WORKING DRAFT
DEFINITION AC023596
ACCESSION AC023596
VERSION AC023596
KEYWORDS HTG; HTGS; PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 177341)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alshrocks,S.L., Amaralunga,H.C., Are,C., Banks,T., Barbaria,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burke,P., Burrell,C., Burrell,K.L., Byrd,N.C., Caron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Covle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flaegge,N., Ford,J.,

```

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

```

Poster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homsli,F., Howard,S., Huber,J., Huliyil,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudan,C., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Licharge,O., Lien,C., Lucier,R., Luna,R.,
Louisgeed,H., Lozano,R.J., Lu,X., Lucier,R., Martindale,A.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabadi,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Mckenko,S.,
Ogih,M., Okunoda,G., Oragunye,N., Oyedele,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Py,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojudoan,I., Rolfe,M.,
Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoshitani,N.,
Slason,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Swack,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalona,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Wesczyk,R., Wooden,S.,
Watlington,S., Williams,G., Williams,A., Wleczky,R., Woodson,D.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 177341)
Worley,K.C.
Direct Submission
Submitted (16-FEB-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 3, 2000 this sequence version replaced gi:10304992.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: RP11-699J21
Center clone name: HAFO
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; 50% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 166578 bases at least Q40
Consensus quality: 166534 bases at least Q30
Estimated insert size: 171835; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 4.2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 32477: contig of 32477 bp in length
* 32478 32577: gap of unknown length
* 32578 51994: contig of 19417 bp in length
* 51995 52094: gap of unknown length
* 52095 71484: contig of 19390 bp in length
* 71485 71585: gap of unknown length
* 84749: contig of 13165 bp in length

```



```

RES
source
    Location/Qualifiers
    1. .139104
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="17"
    /map="17"
    /clone="RP11-377K21"
    /clone_1lb="RPCT-11 Human Male BAC"
    1. .237
misc_feature
    /note="assembly-fragment
clone_end:17
vector_size:left"
338. .1699
misc_feature
    /note="assembly-fragment"
1800. .2964
misc_feature
    /note="assembly-fragment"
3065. .4681
misc_feature
    /note="assembly-fragment"
4782. .6187
misc_feature
    /note="assembly-fragment"
6288. .8839
misc_feature
    /note="assembly-fragment"
8940. .11261
misc_feature
    /note="assembly-fragment"
11362. .13371
misc_feature
    /note="assembly-fragment"
13472. .15817
misc_feature
    /note="assembly-fragment"
15918. .18313
misc_feature
    /note="assembly-fragment"
18414. .20632
misc_feature
    /note="assembly-fragment"
20733. .23241
misc_feature
    /note="assembly-fragment"

```

Query Match	10.4%	Score 47.2	DB 67	Length 139104
Best Local Similarity	48.2%	Pred No. 3.3		
Matches 133	Conservative	0	Mismatches 143	Indels 0
				Gaps 0
Qy 169	tgcaaaccttcttcacggaatbatltaggtltaagtcacgaactggtcctactaagcttt	228		
Db 11206	TATATTAATATTAATAAATGTAAATATTAATAATTAATAATTAATAATTAATAATTA	11447		
Qy 229	atccacacttaatttccaccactttgttgaagaatacacttcacgctccctccactat	288		
Db 11146	TTACATATTTGTATACATTAATATTAATTAATAATTAATAATTAATATTAATTTAT	11087		
Qy 289	atctgtttaatctatgtaaccacaatatctcatgctcaacgccgaatcaactctttat	348		
Db 11086	ATAAATTTATTTATTAATAATAACATATTAATTAATTAATTAATAATTAATTAATAATTA	11027		
Qy 349	gaacagtcttcgctgcgtcgaatacactacaggtttatctacagctgctttatgaaa	408		
Db 11026	TATATATTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATTT	10967		
Qy 409	caaaactacaactcttcaagctctatgctgctttttt	444		
Db 10966	ATATATTAATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATTT	10931		
RESULT 10				
AF031433				
LOCUS	AF031433	1355 bp	DNA	PLN
DEFINITION	Gossypium arboreum tRNA-Thr (ttrn (UGU)) and tRNA-Ieu (trln (UAA)) partial genes, chloroplast genes encoding chloroplast RNA, partial sequence, and trnt (UGU)-trml (UAA) intergenic region.			
ACCESSION	AF031433			

VERSION	AF031433.1	GI:2623897
KEYWORDS	Gossypium arboreum.	
SOURCE	Chloroplast Gossypium arboreum	
ORGANISM	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta	
REFERENCE	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II	
AUTHORS	Malvales; Malvaceae; Gossypium.	
TITLE	1 (bases 1 to 1355)	
JOURNAL	Small, R.L., Ryburn, J.A., Cronn, R.C., Seelanan, T., and Wendel, J.F.	
REFERENCE	The tortoise and the hare: Choosing between noncoding plastome and	
AUTHORS	nuclear adh sequences for phylogeny reconstruction in a recently	
TITLE	diverged plant group	
JOURNAL	Am. J. Bot. 85 (9), 1301-1315 (1998)	
REFERENCE	2 (bases 1 to 1355)	
AUTHORS	Small, R.L., Ryburn, J.A., Cronn, R.C., Seelanan, T., and Wendel, J.F.	
TITLE	Direct Submission	
JOURNAL	Submitted (27-OCT-1997) Botany, Iowa State University, 353 Bessey	
FEATURES	Hall, Ames, IA 50011, USA	
SOURCE	Location/Qualifiers	
	1..1355	
	/organism="Gossypium arboreum"	
	/organism="Gossypium arboreum"	
	/organism="plastid:chloroplast"	
	/db_xref="taxon:29729"	
trna	complement(<1..20)	
	/gene="trnT (UGU)"	
gene	/product="tRNA-Thr"	
	complement(<1..20)	
misc_feature	/gene="trnT (UGU)"	
	21..1353	
	/note="intergenic spacer"	
trna	1354..>1355	
	/gene="trnL (UAA)"	
gene	/product="tRNA-Leu"	
	1354..>1355	
	/gene="trnL (UAA)"	
BASE COUNT	554 a	139 c 175 g 487 t
ORIGIN		
Query Match	10.4%; Score 47; DB 12; Length 1355;	
Best Local Similarity	48.3%; Pred. No. 4.8;	
Matches 131; Conservative	0; Mismatches 140; Indels 0; Gaps 0;	
OY	107	acataattgagtcattcaccacacctgttactggaatagttgtggaatgtgattgattcatag 166
DB	185	ATTAACCTTAATATTAATATTAATACATATTTTATTAATTAATAGTAGATTCATCTAATATTC 244
OY	167	agtgtcaaaccttcttcataatgaataatagttgattgaattacaactggtctactaagctt 226
DB	245	AAGCTAATGTTATAGAAACCTAATTAATTCGATTTAAATTTTAATATATATATTTTA 304
OY	227	ttaattcaatcctaattttaccacactttgttaagaatatatactcttccatctccatc 286
DB	305	TTATTTATCTATATATATTTATATATTTAATATATATTTTAATATATCTTTTATATCTAGTT 364
OY	287	atatactgttlaactaactgtaacacaatattcactgttcacacacgaatcaactcttla 346
DB	365	ATATTTCTTTTATATTAAT 424
OY	347	ctgaacaatgctctgtgctgtcattacata 377
DB	425	TAGTTATTTTATATTTTATATTTTATATTTTATATTTTATATTTTATATTTTATATTTTACA 455
RESULT 11		
AL359262/c		
LOCUS	162763 bp	DNA
DEFINITION	Human DNA sequence from clone RP11-334013 on chromosome 13,	
ACCESSION	AL359262	complete sequence.
VERSION	AL359262.9	GI:12964305
KEYWORDS	HTG.	
SOURCE	human.	

ORGANISM Homo sapiens
Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 162763)
AUTHORS Tracey/A.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquey@sanger.ac.uk
Requests: clonerequests@sanger.ac.uk
On Feb 19, 2001 this sequence version replaced g1:12718063.
During sequence assembly data is computed from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; information on the WORMPEP
database can be found at:
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 13, constructed by the Sanger Centre Chromosome 13
Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/MGP/Chr13>
RP11-334013 is from the library RPC1-11.2 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBAC3.6
This sequence is the entire insert of clone RP11-334013 The true
left end of clone RP11-105A24 is at 125885 in this sequence. The
true right end of clone RP11-522P22 is at 10371 in this sequence.

FEATURES
source
1..162763
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-334013"
/clone_lib="RPC1-11.2"
134..171
/note="MER66B repeat: matches 405..442 of consensus"
348..508
/note="LTR43 repeat: matches 132..296 of consensus"
660..1735
/note="T1 repeat: matches 3870..4955 of consensus"
1741..2133
/note="LTR2CB repeat: matches 1..423 of consensus"
2139..2812
/note="L1MA8 repeat: matches 5158..5850 of consensus"
2814..3014
/note="L1MB2 repeat: matches 5952..6163 of consensus"
3526..3684
/note="MER5A repeat: matches 30..189 of consensus"
3758..3916
/note="LTR1D repeat: matches 352..499 of consensus"
3917..4221
/note="V1usx repeat: matches 1..307 of consensus"
4222..4330
/note="LTR1D repeat: matches 198..352 of consensus"
4435..4487
/note="8 copies 4 mer acac 90% conserved"
7714..7944
/note="Mir repeat: matches 2..249 of consensus"
8322..8866
/note="MER41A repeat: matches 3..554 of consensus"
10653..10836

```
repeat_region /note="92 copies 2 mer tt 59% conserved"
10808. 10863
repeat_region /note="14 copies 4 mer cctc 71% conserved"
11705. 12067
repeat_region /note="MER47A repeat: matches 3. 366 of consensus"
13870. 14231
repeat_region /note="MER67D repeat: matches 1. 340 of consensus"
18418. 18688
repeat_region /note="AluDb repeat: matches 5. 283 of consensus"
22171. 22452
repeat_region /note="L1MC3 repeat: matches 7429. 7735 of consensus"
23271. 23298
repeat_region /note="7 copies 4 mer aaca 92% conserved"
24848. 25683
repeat_region /note="L1 repeat: matches 2350. 3164 of consensus"
25785. 26264
repeat_region /note="240 copies 2 mer ta 55% conserved"
26320
misc_feature /note="Weak data"
26490. 26633
repeat_region /note="72 copies 2 mer ta 60% conserved"
26511. 26634
repeat_region /note="31 copies 4 mer atat 61% conserved"
26634. 26679
repeat_region /note="23 copies 2 mer at 76% conserved"
26796. 26847
repeat_region /note="26 copies 2 mer ga 76% conserved"
27188. 27379
repeat_region /note="L1M4 repeat: matches 4614. 4812 of consensus"
27730. 28333
repeat_region /note="L1M3A repeat: matches 5298. 5951 of consensus"
30264. 30320
repeat_region /note="L2 repeat: matches 2606. 2661 of consensus"
30627. 31083
repeat_region /note="L2 repeat: matches 2020. 2527 of consensus"
31378. 31662
repeat_region /note="AluSg repeat: matches 23. 307 of consensus"
32773. 32812
repeat_region /note="L2 repeat: matches 2238. 2277 of consensus"
33003. 33130
repeat_region /note="L2 repeat: matches 2538. 2674 of consensus"
33237. 33784
repeat_region /note="MLT2A repeat: matches 1. 453 of consensus"
34490. 34696
repeat_region /note="L1M1 repeat: matches 5407. 5613 of consensus"
34698. 34925
repeat_region /note="L1 repeat: matches 5168. 5403 of consensus"
34925. 35042
repeat_region /note="L1 repeat: matches 4513. 4630 of consensus"
35045. 35531
repeat_region /note="L1 repeat: matches 3187. 3672 of consensus"
35531. 37715
repeat_region /note="L1M2 repeat: matches -673. 1545 of consensus"
43406. 43672
repeat_region /note="AluSg1 repeat: matches 1. 296 of consensus"
43830. 44106
repeat_region /note="L1M2 repeat: matches 16. 304 of consensus"
44438. 44717
repeat_region /note="L2 repeat: matches 2405. 2703 of consensus"
44762. 45209
repeat_region /note="L2 repeat: matches 1813. 2289 of consensus"
45279. 45583
repeat_region /note="L2 repeat: matches 1229. 1554 of consensus"
45639. 45797
repeat_region /note="MIR repeat: matches 81. 248 of consensus"
47207. 48663
repeat_region /note="L1M1 repeat: matches 4500. 6117 of consensus"
48811. 48902
repeat_region /note="MIR repeat: matches 70. 163 of consensus"
49688. 49982
repeat_region /note="AluSg repeat: matches 2. 302 of consensus"
50030. 50123
repeat_region /note="47 copies 2 mer tc 76% conserved"

repeat_region 50039. 50118
/note="20 copies 4 mer cctc 82% conserved"
repeat_region 52618. 52916
/note="L2 repeat: matches 179. 504 of consensus"
repeat_region 52975. 53182
/note="L2 repeat: matches 645. 858 of consensus"
repeat_region 53309. 53359
/note="MER53 repeat: matches 139. 189 of consensus"
repeat_region 53423. 53515
/note="MER53 repeat: matches 29. 122 of consensus"
repeat_region 53517. 53634
/note="59 copies 2 mer tt 61% conserved"
repeat_region 53714. 54065
/note="L2 repeat: matches 1172. 1561 of consensus"
repeat_region 55215. 55388
/note="MER20 repeat: matches 1. 163 of consensus"
repeat_region 56456. 56696
/note="AluDb repeat: matches 1. 235 of consensus"
repeat_region 57571. 57918
/note="L2 repeat: matches 1996. 2380 of consensus"
repeat_region 58095. 58201
/note="L2 repeat: matches 2634. 2750 of consensus"
repeat_region 58275. 58312
/note="19 copies 2 mer tt 78% conserved"
repeat_region 58353. 58412
/note="15 copies 4 mer gaaa 75% conserved"
repeat_region 59013. 59610
/note="MER4B repeat: matches 1. 575 of consensus"
repeat_region 60115. 60597
/note="MLT1D repeat: matches 3. 501 of consensus"
repeat_region 60600. 60946
/note="L2 repeat: matches 525. 873 of consensus"
repeat_region 61270. 61299
/note="15 copies 2 mer aa 93% conserved"
repeat_region 61349. 61723
/note="MLT1B repeat: matches 8. 390 of consensus"
repeat_region 63420. 63548
/note="MIR repeat: matches 119. 254 of consensus"
repeat_region 65053. 65557
/note="MER84 repeat: matches 2. 508 of consensus"
repeat_region 65558. 65629
/note="MER84 repeat: matches 1. 74 of consensus"
repeat_region 66347. 66822
/note="MER52A repeat: matches 545. 746 of consensus"
repeat_region 66868. 67056
/note="LTR28 repeat: matches 1. 174 of consensus"
repeat_region 67112. 67208

Query Match 10.1%; Score 45.8; DB 90; Length 162763;
Best Local Similarity 47.9%; Pred. No. 6.1;
Matches 195; Conservative 0; Mismatches 207; Indels 5; Gaps 2;
```

	TITLE	JOURNAL	COMMENT
OY	373 atatactacggttattcctaagtcgtttatgaagaacaactcaaa	419 	
Dn	ATATAAATGCAAAATATTATTTTAAGTGAATATACAAAATATATTAAT	25858	
RESULT	12		
LOCUS	ACO16462		
DEFINITION	Homo sapiens clone RP11-1001B, WORKING DRAFT SEQUENCE, 13 unordered pieces.	HTG -09-JUN-2000	
ACCSSION	ACO16462		
VERSION	ACO16462.5 GI:8389550		
KEYWORDS	HTG; HTGS_PHASEI; HTGS_DRAFT;		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 170274) Birren,B., Linton,L., Nusbaum,C. and Lander,E. Homo sapiens, clone Rp11-1001B Unpublished 2 (bases 1 to 170274)		
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,I., Boukhalter,B., Brown,A., Castle,A., Collange,M., Collins,S., Collymore,A., Cooke,P., DeRellano,K., Dewar,K., Domilo,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest.C., Funke.R., Gage,D., Galagan.J., Gardyna.S., Grant.G., Hagos.B., Headford.A., Horton.L., Howland,J.C., Johnson,R., Jones.C., Kann,L., Karatas,A., Klein,J., Lehocaky,J., Lieu,C., Locke.K., Macdonald.P., Marquis.N., McEwan,P., McGurk,A., McKernan,K., McLoughlin,J., Meldrum,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stoianovic,N., Subramanian,A., Talams,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman.D., Ye.W.J., Zimmer,A. and Zody.M.		
JOURNAL	Direct Submission		
COMMENT	Submitted (30-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Jun 9, 2000 this sequence version replaced gi:8137565. All repeats were identified using RepeatMasker: Smtl, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html Center: Whitehead Institute/ MIT Center for Genome Research Web site: http://www.seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu Project Information Center project name: L3125 Center clone name: LO.O.18 Summary Statistics Sequencing Vector: M13; M77815; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Assembly program: Phrap; Version 0.960731 Consensus quality: 161033 bases at least Q40 Consensus quality: 165923 bases at least Q30 Consensus quality: 167907 bases at least Q20 Insert size: 170000; agarose-if Quality coverage: 169074; sum-of-contigs Quality coverage: 4.5 in Q20 bases; agarose-if Quality coverage: 4.6 in Q20 bases; sum-of-contigs NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will		

Query Match	10.1%	Score 45.8	DB 64	Length 170274
Best Local Similarity	47.9%	Pred. No. 6.1		
Matches 195	Conservative	0	Mismatches 207	Indels 5; Gaps 2;

Source	Location/Qualifiers
misc_feature	1. 1270: contig of 1270 bp in length
misc_feature	1271 1370: gap of 100 bp
misc_feature	1371 7291: contig of 5921 bp in length
misc_feature	7292 7391: gap of 100 bp
misc_feature	7392 14101: contig of 6710 bp in length
misc_feature	14102 14201: gap of 100 bp
misc_feature	14202 20367: contig of 6166 bp in length
misc_feature	20368 20467: gap of 100 bp
misc_feature	20468 24965: contig of 4498 bp in length
misc_feature	24966 25065: gap of 100 bp
misc_feature	25066 32957: contig of 7792 bp in length
misc_feature	32958 32957: gap of 100 bp
misc_feature	32958 45121: contig of 12164 bp in length
misc_feature	45122 45221: gap of 100 bp
misc_feature	45222 57321: contig of 12000 bp in length
misc_feature	57322 57321: gap of 100 bp
misc_feature	57322 73326: contig of 16005 bp in length
misc_feature	73327 73426: gap of 100 bp
misc_feature	73427 88792: contig of 16366 bp in length
misc_feature	88793 89892: gap of 100 bp
misc_feature	89893 107479: contig of 17587 bp in length
misc_feature	107480 107579: gap of 100 bp
misc_feature	107580 125744: contig of 18165 bp in length
misc_feature	125745 125845: gap of 100 bp
misc_feature	125845 170274: contig of 44430 bp in length.
FEATURES	
source	1. 170274
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="RP11-10018"
	/clone_lib="RPC1-11 Human Male BAC"
	1. 1270
misc_feature	/note="assembly_fragment"
misc_feature	1371. 7291
misc_feature	/note="assembly_fragment"
misc_feature	7392. 14101
misc_feature	/note="assembly_fragment"
misc_feature	14202. 20367
misc_feature	/note="assembly_fragment"
misc_feature	20468. 24965
misc_feature	/note="assembly_fragment"
	clone_end:SP6
	vector_side:right"
misc_feature	25066. 32857
	/note="assembly_fragment"
	clone_end:T7
	vector_side:left"
misc_feature	32958. 45121
misc_feature	/note="assembly_fragment"
misc_feature	45222. 57221
misc_feature	/note="assembly_fragment"
misc_feature	57322. 73326
misc_feature	/note="assembly_fragment"
misc_feature	73427. 88792
misc_feature	/note="assembly_fragment"
misc_feature	88983. 107479
misc_feature	/note="assembly_fragment"
misc_feature	107580. 125744
misc_feature	/note="assembly_fragment"
misc_feature	125845. 170274
misc_feature	/note="assembly_fragment"
BASE COUNT	55400 a 30429 c 30429 g 52813 t 1203 others
ORIGIN	


```

misc_feature      /note="assembly-fragment"
                  154493..178068
                  /note="assembly-fragment"
misc_feature      178169..186695
                  /note="assembly-fragment"
                  clone_end:17
                  vector_side:right"
BASE COUNT      58821 a 32954 c 32688 g 61023 t 1209 others
ORIGIN

Query Match      10.1%; Score 45.8; DB 63; Length 186695;
Best Local Similarity 47.9%; Pred. No. 6.1;
Matches 195; Conservative 0; Mismatches 207; Indels 5; Gaps 2;

OY 18 tcaagacgtacatgaattgactgattacattaccagaagaatgttaatacacc 77
    || || || || || || || || || || || || || || || || || || ||
DB 25256 TATATATATCAATATATTTATATATATATATATATATATATATATATAT 25315

OY 78 ctggtagacatgttggaggatttactccacaataatgagtcacatccactgtact 137
    || || || || || || || || || || || || || || || || || || ||
DB 25316 TATTATATATATTTATATATATATATATATATATATATATATATATATAT 25375

OY 138 ggaatagctgtcggaattgtagtcaagtgcaacttcttcaatggaatattag 197
    || || || || || || || || || || || || || || || || || || ||
DB 25376 ACAATATATATATATATATATATATATATATATATATATATATATATATA 25435

OY 198 tttaagttacaactgcttactaagcttattacacatccttaatttccactttt 257
    || || || || || || || || || || || || || || || || || || ||
DB 25436 TATTAATATATATATATATATATATATATATATATATATATATATATATAT 25495

OY 258 taagaatataccttctcagctctccacatatac---tgtttaactatgtaccac 313
    || || || || || || || || || || || || || || || || || || ||
DB 25496 ATATATATATATATATATATATATATATATATATATATATATATATATATAC 25555

OY 314 aatattatgtccaacccaagaatcaatcttcttcttcttcttcttcttcttct 372
    || || || || || || || || || || || || || || || || || || ||
DB 25556 ATTATATATCAATATATATATATATATATATATATATATATATATATATAT 25615

OY 373 atatactagcttcttctcagcttcttcttcttcttcttcttcttcttcttct 419
    || || || || || || || || || || || || || || || || || || ||
DB 25616 ATATATATCAATATATATATATATATATATATATATATATATATATATAT 25662

RESULT 14
AL353783      182232 bp      DNA      HTG      06-MAR-2001
LOCUS      Homo sapiens chromosome 9 clone RP11-27103, *** SEQUENCING IN
DEFINITION
ACCESSION      AL353783
VERSION      AL353783.6 GI:13273652
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Plumb.B.
Direct Submission
Submitted (01-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
On Mar 11, 2001 this sequence version replaced gi:1324872.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba27103
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; 108752; 100% of reads

```

```

Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 180394 bases at least Q40
Consensus quality: 181011 bases at least Q30
Consensus quality: 181436 bases at least Q20
Insert size: 181832; sum-of-contigs
Insert size: 173385; 3.0% error; agarose-1p
Quality coverage: 8.57x in Q20 bases; sum-of-contigs quality
coverage: 9.10x in Q20 bases; agarose-1p
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 47401: contig of 47401 bp in length
* 47402 47501: gap of 100 bp
* 47502 82489: contig of 34988 bp in length
* 82490 82589: gap of 100 bp
* 82590 150067: contig of 67478 bp in length
* 150068 150167: gap of 100 bp
* 150168 155571: contig of 5404 bp in length
* 155572 155671: gap of 100 bp
* 155672 182232: contig of 26561 bp in length.
FEATURES
Source
1..182232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-27103"
/clone_id="RCR1-11.1"
misc_feature
1..47401
/note="assembly-fragment: 00307
clone_end:SP6
vector_side:left"
47502..82489
/note="assembly-fragment: 00972
fragment_chain:1"
82590..150067
/note="assembly-fragment: 03273
fragment_chain:1"
150168..155571
/note="assembly-fragment: 02310"
155672..182232
/note="assembly-fragment: 02328"
BASE COUNT      59904 a 33351 c 30550 g 58027 t 400 others
ORIGIN

Query Match      10.1%; Score 45.6; DB 79; Length 182232;
Best Local Similarity 47.8%; Pred. No. 6.6;
Matches 132; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

OY 169 tctcaacttcttcaatgaatattaggttaagttacacactgcttcaagcttt 228
    || || || || || || || || || || || || || || || || || || ||
DB 156077 TATATATATATATATATATATATATATATATATATATATATATATATATA 156136

OY 229 atccacatcttaatttaccacttctgttgaataataccttcaagctctccacat 288
    || || || || || || || || || || || || || || || || || || ||
DB 156137 TTACATATTTTGATACATTTATATATATATATATATATATATATATATAT 156196

OY 289 atctgttaatactatgttaaccaataatcaatgcaacccaagaatcaatccttact 348
    || || || || || || || || || || || || || || || || || || ||
DB 156197 ATAAATTTTATATATATATATATATATATATATATATATATATATATAT 156256

OY 349 gacatgttctggttgcataacatactacaggtttacagtgcttcttataagaa 408
    || || || || || || || || || || || || || || || || || || ||
DB 156257 TATATATATATATATATATATATATATATATATATATATATATATATAT 156316

OY 409 caaactcaacttcttaagttctatgtgtttt 444
    || || || || || || || || || || || || || || || || || || ||

```


us-09-784-340-3_copy_18322_18774.rge

Page 17

Search completed: August 27, 2001, 17:57:00
Job time: 8893 sec

THIS PAGE BLANK (USPTO)